AD_____

GRANT NUMBER DAMD17-96-1-6084

TITLE: Metastic Regulation of Differential Splicing of CD44

PRINCIPAL INVESTIGATOR: Susan M. Berget, Ph.D.

CONTRACTING ORGANIZATION: Baylor College of Medicine
Houston, Texas 77030-3498

REPORT DATE: August 1997

TYPE OF REPORT: Annual

PREPARED FOR: Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for public release;
distribution unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

19971218 023

DTIC QUALITY INSPECTED 4

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1997	Annual (15 Jul 96 - 14 Jul 97)

4. TITLE AND SUBTITLE Metastic Regulation of Differential Splicing on CD44

5. FUNDING NUMBERS
DAMD17-96-1-6084

6. AUTHOR(S)
Susan M. Berget, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Baylor College of Medicine
Houston, Texas 77030-3498

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Commander
U.S. Army Medical Research and Materiel Command
Fort Detrick, Frederick, Maryland 21702-5012

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200*

CD44 is a multifunctional adhesion molecule that can undergo extensive alternative RNA splicing to generate multiple isoforms bearing different extracellular domains, a sub-set of which have been correlated with metastasis in breast cancer. Three such exons, variable exons 5, 6, and 7, were studied for inclusion in RNA from normal and neoplastic tissues from the mouse, using a recently described system for growth and transplantation of mammary epithelial cell lines. Variable exon 6, but not variable exons 1-5, or 7, was included in CD44 mRNA produced in normal mature mammary gland, indicating that splicing of CD44 is restricted in normal adult mammary tissue. Pregnant or lactating gland expressed variable exon 6, and, at lower levels, variable exons 5 and 7. Patterns of variable exon splicing in the TM2L hyperplasia and the TM40D preneoplasia were similar to that of tissue from pregnant animals, suggesting induction of a normal developmental pattern during hyperplastic growth. Considerable alternative splicing of multiple exons from the v1-v7 region of CD44 was observed in tumors and metastases. These results suggest that hyperplasias and preneoplasias access an alternative processing pattern that is turned on during normal early pregnancy and that this pattern is further accentuated during tumorogenesis.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	33
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

SWB Where copyrighted material is quoted, permission has been obtained to use such material.

SWB Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

SWB In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and Use of Laboratory Animals of the Institute of Laboratory Resources, National Research Council (NIH Publication No. 86-23, Revised 1985).

____ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

SWB In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

SWB In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

SWB In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

Susan M. Bright Aug 11, 1997

PI - Signature Date

(4) TABLE OF CONTENT

FRONT COVER	1
STANDARD FORM (SF) 298	2
FOREWORD	3
TABLE OF CONTENTS	4
INTRODUCTION	5
BODY	6
CONCLUSIONS	19
REFERENCES	20
APPENDICES	24

(5) INTRODUCTION

An early step during tumor cell metastasis is dissemination from the primary site. After penetration of the extracellular matrix (ECM), tumor cells invade the vascular system (blood and lymph) and establish secondary foci through interactions with epithelial and endothelial cells at organ-specific sites. In response to angiogenesis and growth factors, such re-located cells build tumor-specific tissue and organ metastases. The capacity of tumor cells to penetrate the ECM, as well as their subsequent cell-cell interactions are mediated by different families of cell adhesion molecules. One such adhesion protein is the multifunctional transmembrane glycoprotein CD44 (1), expressed in a variety of tissues. CD44 pre-mRNA undergoes extensive alternative RNA splicing to create variants with exon-specific extracellular domains (2,3). Expression of a sub-set of these variable domains has been correlated with the capacity of tumor cells to metastasize and progression of disease in colon, breast, brain, and ovarian cancer (4-6). Consequently, the state of splicing of CD44 has been used as a marker for metastasis.

Alternative splicing of CD44 to express a set of variable exons has also been shown to be determinative for metastasis in mice. Non-metastatic tumor cells expressing a standard form of CD44 lacking the variable exons were transfected with a cDNA coding for alternatively spliced CD44 containing variable exon 6. These transfected cells were metastatic when reimplanted in syngenic animals (7). In addition, monoclonal antibodies directed against the metastasis-associated isoforms of CD44 block metastasis in a mouse model (8). These observations strongly implicate CD44 alternative splicing as an important parameter in metastasis of at least some tumors.

Although multiple studies have appeared correlating the state of splicing of CD44 in various human tumors, little work has assessed the relationship of CD44 alternative splicing to early events in disease progression. Here we have used an *in situ* mouse model system (9-11) to investigate the state of CD44 splicing during normal breast development, hyperplasias, preneoplasias, tumors, and metastases. This system takes advantage of established mouse mammary epithelial cell lines (MMEL cell lines) and subsequent alveolar hyperplastic outgrowths and transplanted ductal outgrowths of these lines generated by in vivo transplantation into the mouse mammary fat pad. We find that limited alternative splicing of CD44 to include metastatic exon 6 occurs in adult virgin breast. Increased alternative splicing to include metastatic variable exons 5, 6, and 7 occurs in early pregnancy and lactating breast. Hyperplasias and preneoplasias had a pattern of alternative splicing of variable exons 5, 6, and 7, that closely resembled that of pregnant gland. In contrast, all examined tumors had extensive alternative splicing of variable exons 1-7. Progression of disease was also accompanied by increased expression of the arginine and serine-rich family of splicing factors (SR proteins) thought to regulate inclusion of multiple variable exons, including the non-metastatic variable exons of CD44 expressed in epithelial cells.

(6) BODY

Experimental Methods, Assumptions, and Procedures

Tissues. The splicing patterns of the CD44 gene were studied in a mouse model of mammary development (9-11) and tumorigenesis in normal tissue, pregnant mice,

lactating glands, preneoplastic lesions, 8 primary breast cancers and metastasis of different locations (liver and lung). Development of this system has been described elsewhere (9-11). In general, the preneoplastic lesions were raised from hyperplastic TM outgrowth cell lines (HOG) after transplantation in cleared mammary fat pads of 3 week old syngenic BALB/c mice. The outgrowths were removed after 8-12 weeks. Some lesions were left for further tumor development *in vivo*. The tumor incidence was individual for different HOG's as well as the ability of this tumors to metastasize. After removal of the fat pads, filled with the HOG's preneoplastic lesions as well as removal of primary adenocarcinomas or metastases, tissues were frozen at –.80°C for further analysis.

RNA isolation and RT/PCR analysis.. Tissues were minced on dry ice in little pieces before they were treated with a tissue homogenizer (Polytron) in TRIzol solution (Gibco) at 1ml per 100 mg of tissue for 3 times 10 sec. After 5 minutes of incubation at .26°C, 0.2 ml of chloroform per 1 ml of TRIzol were added and the solution was shaken vigorously, followed by incubation for 5 minutes. After centrifugation for 15 min at 4oc at 12,000 x g the aqueous phase, containing the total RNA, was precipitated with 0.5 ml of isopropyl alcohol per 1 ml of TRIzol for 10 minutes and cenfrifuged for 10 min at 12,000 x g. The pellets were rinsed with 75% ethanol (1ml of ethanol per 1 ml of TRIzol reagent used for the initial homogenization), vortexed and centrifuged at .40°C for 5 min at 7.500 x g and finally dissolved in RNAse free water and stored at –80°C for further analysis.

Five micrograms of RNA were used for cDNA synthesis using M-MuLV reverse transcriptase (Perkin Elmer) and oligo-dT primers followed by PCR using 3' primers specific for individual alternative exons (CD44v5, v6 or v7) as well as for the CD44 standard form (CD44 std, from exon 16) and a specific ATP-labeled 5' prime derived from exon 5, respectively. The PCR conditions for CD44 amplification were as follows:

25 cycles of 94°C for 1 min, 59°C for 1 min and 72°C for 1.5 min. High cycle amplification used 35 cycles.

Gel electrophoresis of 35 µl of the 100 µl PCR reaction was performed in 6% denaturing acrylamide gels at 40-50 mA. Markers were. pBR322/Hpa II with bands of 622, 527, 404, 309, 242, 238, 217, 201, 190, 180, 160, 147, 122, 110, 90, 76 bp.

Protein Isolation and Western Blot. Proteins were isolated from the interphase and phenol phase from the initial homogenate after precipitation of the DNA with 0.3 ml 100% of ethanol per 1 ml of TRIzol used for the initial homogenization. Proteins were precipitated with 1.5 ml isopropyl alcohol per 1 ml TRIzol used for initial homogenization. Protein pellets were washed three times in 0.3M guanidine hydrochloride in 95% ethanol for 20 minutes. Pellets were resuspended in 1 % SDS solution and incubated at .50°C for complete dissolution. Insoluble material was removed by centrifugation at 10000 x g for 10 minutes.

Gel electrophoresis of 10 ug of total protein was performed using a 10% SDS-Page gel at 100 Volts. Afterwards the gels were electro-blotted on a PVDF transfer membran (PolyScreen, Dupont) at 100 Volts for 2.5 hours at .4°C. After blocking in 5% Blotto/PBST membranes were incubated with specific IgM mAb 104 (kindly provided by M. Roth, Fred Hutcheson Cancer Cneter) (1:5) for S/R protein detection. The Western blots were stained by chemi-luminescence using a horseradish-peroxidase labeled goat-anti-mouse IgM antibody.

Results and Discussion

CD44 standard is expressed throughout normal murine breast development and neoplasia. To correlate CD44 patterns of splicing with changes associated with normal

and neoplastic breast development, we examined the types of CD44 mRNA present in normal murine breast tissue, hyperplasias, neoplasias, tumors, and metastases of these tumors. All preneoplastic and neoplastic tissues were derived from an *in situ* system comprised of outgrowths of re-implanted breast epithelial cell lines (9-11). One hyperplasia, TM2L produces tumors at very low frequency (10). A second preneoplasia, TM40D, has a higher incidence of tumor appearance and these tumors metastasize (D. Medina, personal communication). In addition we studied multiple tumors produced from various cell lines and metastases thereof. We concentrated on determining the splicing pattern of the three alternative exons, v5, v6, and v7, for which inclusion has been correlated with metastasis.

We used low cycle RT/PCR to examine the splicing pattern of CD44 RNA. CD44 standard was detected using PCR primers complementary to constitutive exons 5 and 16. These two exons border the central region of the CD44 gene encoding the ten alternative exons (Fig. 1). Thus, RT/PCR with these primers should give an indication of the overall amounts of CD44 RNA present in a tissue and some indication of the extent of alternative splicing. Due to the presence of ten alternative exons between exons 5 and 16, however, amplification patterns with these primers can yield complicated patterns depending upon the number of exons included in the final mRNA. Therefore, we also used variable exon-specific primers to detect the presence of exons v5, v6, and v7 in RNA from tissues.

Amplification of RNA with primers specific for exons 5 and 16 indicated that mRNA coding for CD44 standard, lacking the variable exons, was expressed in all tissues examined (Figure 2). The mature virgin gland, hyperplasias, and preneoplasias contained

less overall CD44 mRNA than lactating or pregnant tissue, or tumors, suggesting that transcription of CD44 is increased during pregnancy and in tumors.

We wished to be able to distinguish CD44-specific affects on alternative splicing associated with neoplastic changes from generic affects on splicing activity associated with accelerated growth rate. Therefore, we also examined the amounts and spliced forms of an RNA that undergoes stochastic alternative splicing in most tissues. For this purpose we used the hnRNP A1 mRNA (12,13). This ubiquitously-transcribed RNA is alternatively processed in most tissues in which the RNA is expressed to produce two mRNA isoforms which alternatively include exon 7B. Examination of the ability of exon 7B to be included in different tissues affords a first approximation for the level of RNA processing factors in a given tissue. Relatively equal levels of hnRNP A1 were produced in all tissues examined (Fig. 2). With the exception of mature mammary gland, equal levels of exon 7B inclusion were observed, indicating that the examined breast tissues have relatively normal levels of constitutive splicing factors. Mature gland had a different pattern of hnRNP A1 expression in that the percentage of exon 7B inclusion was noticeably higher than in other tissues. This result indicates that none of the tissues under study has a noticeable lack of processing factors, because no tissue lacked inclusion of exon 7B. The presence of high inclusion of exon 7B in mature breast gland is unusual, and suggests the presence of special factors required for exon 7B recognition in this tissue.

Multiple CD44 variable exons are included in tumor RNA. Low cycle RT/PCR with primers specific for exons 5 and 16 indicated low levels of inclusion of variable exons in

normal breast tissues, as revealed by the absence of higher molecular weight bands in Figure 2. In contrast, tumor RNA revealed amplification of RNAs containing multiple variable exons. The molecular weights of the observed bands indicated that multiple exons had been incorporated in a single RNA. Thus, despite the fact that splicing of the marker RNA, hnRNP A1, did not increase in tumor tissue, alternative splicing of CD44 increased markedly, suggesting induction of factors specific for CD44 splicing, rather than a general increase in generic splicing factors, during tumorigenesis.

Metastatic exons in normal development. To better examine variable exon splicing, RNA from tissue samples was analyzed by RT/PCR using variable exon-specific primers (Fig. 3). For this purpose reverse transcription and amplification used primers specific for variable exons v5, v6, and v7, coupled with a PCR primer for constitutive exon 5. Such an amplification strategy permits examination of the inclusion of all alternative exons in the interval v1–v7. It does not address inclusion of variable exons v8-v10. Both low and high cycle amplifications were performed.

In mature gland, little inclusion of v5 or v7 could be detected following 25 cycles of amplification (Fig. 3A, and 3C). Higher cycles revealed very low levels of inclusion of variable exon 7 but little to no variable exon 5 (Fig. 4). Thus, mature gland produced little to no mRNA including the so-called metastatic exons, v5 or v7. Mature gland did produce a band of 210 nucleotides, diagnostic of an RNA containing v6 but not v1-v5. Thus, one metastatic exon, v6, is included in mature virgin non-neoplastic breast tissue.

Both pregnant and lactating gland, in contrast, showed evidence of inclusion of the three metastatic exons when their presence was investigated using both low and high

cycle amplification with exon-specific primers (Fig. 3 and 4). It should be noted that the exons in the interval v1-v7 are all very similar in size, such that inclusion of each exon causes an increase in molecular weight of the observed PCR product by approximately 100 nucleotides. Amplification with a v5-specific primer indicated two RNA species, one with one exon from the interval v1-v5 and one with two exons (PCR products of 208 and 323-335 nucleotides, respectively). The former RNA should include v5 and no other upstream variable exons; the latter should contain v5 plus one other exon from the v1-v4 region. Amplification with a v6-specific primer indicated the production of an RNA containing v6 and no other upstream exon (PCR product of 210 nucleotides). Amplification with a v7-specific primer indicated two species; one with just v7 and one with v7 plus one upstream exon (PCR products of 209 and 324-336 nucleotides, respectively). Amplification products are presently being sequenced to determine the identities of the observed species.

Higher cycle amplifications provided a better view of the RNA species containing v5 and v7 (Fig. 4). V5-containing RNAs with one, two, three, and four variable exons could be detected indicating that low levels of splicing to include multiple variable exons occurs in pregnant tissue. Similar high cycle analysis of RNAs containing v7 indicated only two species of RNA, suggesting less complexity in the pattern of RNAs containing v7. These RNA species were in less abundance than RNAs containing v6, which could be detected easily with low cycle RT/PCR (Fig. 3B).

Splicing of CD44 variable exons in neoplastic tissue. The two pre-neoplasias TM2L and TM40D revealed inclusion of v5, v6, and v7 in a pattern very similar to that of

pregnant tissue (Fig. 3 and 4). High levels of v6 inclusion, and low but detectable levels of inclusion of v5 and v7 were observed. This observation suggests that the process occurring during production of these outgrowth epithelial cells was accompanied by an RNA processing pattern similar, but not identical, to that induced during pregnancy.

In contrast, primary breast tumors and their metastases demonstrated pronounced increases in alternative splicing of CD44 (Fig. 3). Variable exon 5 was included in RNA products that contained one to five of the variable exons v1-v5, indicating an overall increase in CD44 alternative splicing in tumors. This alteration is in strong contrast to the absence of increased alternative splicing of the marker hnRNP A1 RNA. V6 alternative splicing also increased in tumors. Like v6, v7 was included in a battery of RNAs in tumors, but the pattern of created RNAs was simpler than when v5 was examined. In the case of v7, the most abundant RNA species were those including one to three variable exons from v1-v7.

Variable exon v3 is only expressed in tumors. The pattern of tumor RNA alternative splicing was very complex, suggesting that many of the variable exons in the region v1-v7 were becoming included in RNA. To examine this issue directly, amplification of RNA was performed using variable exon 3-specific primers (Fig. 5). Amplification of RNA from mature gland, pregnant gland, the hyperplasia TM2L, and the preneoplasia TM40D indicated the presence of no RNA containing variable exon 3. Tumor RNA, however, indicated roughly equal levels of RNA containing one exon, two exons, and three exons from the internal v1-v3. Thus, considerable alternative splicing of CD44

variable exons was induced during tumorigenesis to produce RNAs containing exons not included in RNAs produced in normal breast development.

Alternative splicing of CD44 in cell lines derived from breast epithelial cell outgrowths. The murine model system, employed also has developed cell lines from individual outgrowths. By analysis of various markers, including p53 status, these cell lines appear clonal (11). Thus, these cell lines provide a uniform cell population in which to examine CD44 alternative splicing associated with the preneoplastic state. We examined CD44 alternative splicing in two of these cell lines, TM3 and TM10 (9-11). The goal of this experiment was to determine if the low levels of splicing of the variable exons v5 and v7 observed in preneoplastic tissue reflected a low frequency event in all cells or an event associated with a small sub-population of cells. Variable exons v5 and v7 were included in RNA from these lines to produce an RT/PCR pattern very similar to that observed with the tissues depicted in Figure 3 (data not shown). Therefore, we interpret the low levels of exons 5 and 7 revealed in these tissues to be the result from low levels of inclusion of these exons in all cells, instead of a high level of inclusion in a small population of different cells in the tissue.

Tumorgenesis increases expression of SR proteins. SR proteins are a large family of RNA binding proteins implicated as involved in alternative splicing (14-16). These proteins contain extensive regions rich in arginine and serine (the SR domain) in addition to their RNA binding domains. The SR domain is necessary for full activity of the proteins and is thought to represent a protein-protein interaction domain. Five to six

major SR proteins are recognized by a monoclonal antibody, mAb104, which recognizes a phosphorylated epitope within the SR domain (16). In HeLa cells, mAb 104 recognizes the major proteins SRp75, SRp55, SRp40, SRp30a, SRp30b, and SRp20 of molecular weights 75, 55, 40, 30, and 20 kDa, respectively. Expression of these major members of the family varies in different tissues (16, 17).

In pregnant and lactating breast tissue, SR protein expression was limited as detected by Western blotting with mAb104 (Fig. 6). Only SRp75 and SRp30b were detected n significant amounts. In contrast, tumor cells expressed the entire spectrum of SR proteins recognized by this antibody. These results suggest that SR protein expression is induced upon tumorigenesis. Considering the wide range of target RNAs thought to be recognized by SR proteins (14-20), such induction should alter the splicing patterns of a wide number of alternatively processed RNAs. We are presently preparing protein from preneoplastic tissues and from cell lines to examine the levels of SR protein expression in hyperplasias and preneoplasias.

The antibody used for the Western Blot in Figure 5 requires phosphorylation of SR proteins for detection (16). Thus, it is formally possible that there are non-phosphoryated SR proteins present in pregnant gland that were not detected in this experiment. We have other antibodies to these proteins in the laboratory and will use these antibodies to determine total SR protein levels.

Cloning and sequencing of the human CD44 gene region containing the variable exons v4 – v7. To understand the mechanism of alternative splicing of the CD44 variable exons and the relationship of that splicing to cancer, we need to be able to create

simple transgenes containing portions of the CD44 gene. As our goal is to study the metastatic exons, our initial transgenes are designed to contain variable exons v4–v7. Although the human gene has been cloned, the sequence of the gene has not been released. Nor are genomic clones available in the U.S. Therefore, we began this portion of the project by obtaining our own human genomic clones of CD44. This portion of the project was aided by a large repository of human genomic clones at Baylor College of Medicine within the Human Genome Center. A genomic library was screened for relevant CD44 sequences through the use of primers specific for intron 9. This short intron within the variable region of interest is the only intronic portion of the CD44 gene for which sequence information has been released. Using this region as probe, we identified two large genomic inserts of 40 kB containing the region of interest.

One sub clone from this region contains a segment of DNA encompassing exons v4 and v5 and their surrounding intron regions. This sub-clone has been entirely sequenced and is being analyzed for known RNA processing sequence motifs. We anticipate from work on other systems that the sequences necessary for inclusion of these exons will reside within each exon and its surrounding intron. We already know that the exons of interest contain sequences similar to known exon enhancers that are the targets for the SR proteins (18, 20). We don't yet know what intronic sequences will be important for alternative exon recognition. When the exon and its surrounding intron is re-cloned into the intron of a heterologous gene and transfected into appropriate cell lines, we should be able to recapitulate in vivo patterns of alternative splicing. Thus, we have made clones in which we have inserted exons v4 and/or v5 as internal exons in a heterologous mini-gene providing a promoter, a first exon, and a last exon and poly(A)

site. These mini-genes have been expressed and spliced in transfected HeLa cells, indicating that a functional gene has been created. We are just starting to learn how to transfect breast epithelial cell lines to begin analysis of the breast factors required to recognize these exons. We are also pressing forward on our analysis and sequencing of genomic clones to identify a similar useful clone containing variable exons v6 and v7 with the goal of establishing a collection of mini-genes containing the regions of interest. Although we had hoped to have all of our cloning and sequencing work done by this time, our cloning efforts have been diluted because of what we consider to be the interesting and exciting analyses of the mouse lines discussed in earlier sections.

Discussion. The observation of alternative inclusion of variable exon v6 in normal breast development indicates that the inclusion of this exon should not be considered as only occurring in cancerous breast. Furthermore, the induction of inclusion of variable exons v5 and v7 in pregnant gland indicates that multiple so-called "metastatic" exons should be considered as exons expressed in CD44 during normal mammary development. It is possible that expression of these exons is used during ductal outgrowth during early pregnancy. CD44 is normally considered as participating in cell-cell and cell-matrix interactions during early development to set up tissues. Thus, changing one of its extra-cellular domains by alternative inclusion of exons v5-v7 could easily alter the protein and its parental cell in a fashion useful to development.

Considerable alternative splicing of CD44 was observed in breast tumors. Multiple variable exons were incorporated into mRNA in these tumors. The patterns of inclusion from tumor to tumor were amazingly similar suggesting generic induction of a

pathway boosting CD44 exon inclusion in breast tumors. Interesting the inclusion of a different alternative exon, that coding for hnRNP A1 exon 7B, was not increased in a similar fashion; suggesting that the boost to CD44 splicing was particular to the gene. We anticipate that CD44 splicing will be responsive to overall levels of SR proteins due to the presence of sequences resembling splicing enhancers within the CD44 variable exons. The increased expression of phosphorylated SR proteins suggests that CD44 exon inclusion levels may increase in tumors because of an increase in the abundance of this family of proteins. We will directly test this hypothesis by transfection of cDNAs coding for SR proteins into the TM cell lines in which CD44 variable exon inclusion is minimal.

We began this study because of published reports that CD44 alternative splicing was correlated with the onset of metastases in a number of human tumors. The role of CD44 alternative splicing in breast cancer metastases, however, has been controversial (21-25). Our study indicates that induction of alternative splicing of CD44 precedes metastasis and, for some exons, precedes the conversion of preneoplasias to tumors. Thus, we suggest that alterations in CD44 alternative splicing are early events in neoplasia. This observation, of course, does not rule out a role for variable exons in the metastatic process.

Recommendations in Relation to the Statement of Work.

The observations reported above with respect to the alterations of CD44 alternative splicing in normal and hyperplastic breast cells are an exciting first view of early events in neoplasia. Immediate planned experiments include an investigation of the

other alternative exons in CD44 with respect to their inclusion in normal, preneoplastic and neoplastic cells. We are especially interested in the epithelial exons v8, v9, and v10 which could play a role in normal breast development or tumor progression.

We also want to further investigate the interesting splicing protein differences that we have observed in tumor cells and extend this analysis to the hyperplasias and neoplasias. We have a battery of antibodies against splicing factors in the laboratory that can be brought to bear on the problem.

And finally, we continue to make mini-genes to study splicing of CD44 in a transfected model system in which we can begin to dissect the cis-acting sequencing controlling this event in early neoplasia.

(7) CONCLUSIONS

CD44 is a multifunctional adhesion molecule that can undergo extensive alternative RNA splicing to generate multiple isoforms bearing different extracellular domains, a sub-set of which have been correlated with metastasis in breast cancer. Three such exons, variable exons 5, 6, and 7, were studied for inclusion in RNA from normal and neoplastic tissues from the mouse, using a recently described system for growth and transplantation of mammary epithelial cells lines. Variable exon 6, but not variable exons 1-5, or 7, was included in CD44 mRNA produced in normal mature mammary gland, indicating that splicing of CD44 is restricted in normal adult mammary tissue. Pregnant or lactating gland expressed variable exon 6, and, at lower levels, variable exons 5 and 7. Patterns of variable exon splicing in the TM2L hyperplasia and the TM40D

preneoplasia were similar to that of tissue from pregnant animals, suggesting induction of a normal developmental pattern during hyperplastic growth. Considerable alternative splicing of multiple exons from the v1-v7 region of CD44 was observed in tumors and metastases. These results suggest that hyperplasias and preneoplasias access an alternative processing pattern that is turned on during normal early pregnancy and that this pattern is further accentuated during tumorigenesis.

To our knowledge this is one of the few reports to look at alternative splicing during early events in neoplasia. The correlation between alternative splicing of CD44 in pregnancy and that observed in hyperplasias and preneoplasias has not been made before and suggests the induction of a normal developmental pathway during hyperplasia. Furthermore, alterations in major alternative splicing factors during early neoplasia have not been examined. The potential for changes in the levels of these proteins to have major impacts on the production of different proteins in tumors is enormous.

(8) REFERENCES

1. Tanabe, K. K., and H. Saya. 1994. The CD44 adhesion molecule and metastasis. Critical Reveiws in Oncogenesis. 5:210-212.

2. Screaton, G. R., M. V. Bell, D. G. Jackson, F. B. Cornelius, U. Gerth, and J. I. Bell. 1992. Genomic structure of DNA encoding the lymphocyte homing receptor CD44 reveals at least 12 alternatively spliced exons. Proc. Natl. Acad. Sci. USA 89:12160-12164.

3. Cooper, D. L., G. Dougherty, H. J. Harn, S. Jackson, E. W. Baptist, J. Byers, A. Datta, G. Phillips, N. R. Isola. 1992. The complex CD44 transcriptional unit: Alternative splicing of three internal exons generates the epithelial form of CD44. Biochem. Biophys. Res. Comm. 182:569-587.

4. Heider, K., M. Hoffman, E. Hors, F. van der Berg, H. Ponta, P. Herrlich, and S. T. Pals. 1993. A human homologue of the rat metastasis-associated variant of CD44 is expressed in colorectal carcinomas and adenomatous polyps. J. Cell. Biol. 120:227-233.

5. Matsumura, Y., and D. Tarin. 1992. Significance of CD44 gene products for cancer diagnosis and disease evaluation. Lancet 240:1053-1058.

6. Kuppner, M. C., E. van Meir, T. Gauthier, M. F. Hamou, and N. D. DeTribolet. 1992. Differential expression of the CD44 molecule in human brain tumors. Int. J. Cancer 50:572-577.

7. Gunthert, U., M. Hofmann, W. Rudy, S. Reber, M. Zoller, I. Haubmsann, S. Matzku, A. Wenzel, H. Ponta, and P. Herrlich. 1991. A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma cells. Cell 65:13024.

8. Sy, M. S., Y. J. Guo, and J. Stamenkovic. 1992. Inhibition of tumor growth in vivo with a soluble CD44-immunoglobulin fusion protein. J. Expt. Med. 176:623-627.

9. Kittrell, C. J. Osborn, and D. Medina. 1992. Development of mammary preneoplasias in vivo from mouse mammary epithelial cell lines in vitro. Cancer Res. 52:1924-1932.

10. Medina, D., F. S. Kittrell, Y.-J. Liu, and M. Schwartz. 1993. Morphological and functional properties of TM preneoplastic mammary outgrowths. Cancer Res. 53:663-667.

11. Ozbun, M.A., D. J. Jerry, F. S. Kittrell, D. Medina, and J. S. Butel. 1993. p53 mutations selected in vivo when mouse mammary epithelial cells form hyperplastic outgrowths are not necessary for establishment of mammary cell lines in vitro. Cancer Res. 53:1646-1652.

12. Buvoli, M., F. Cobianchi, M. G. Bestagno, A. Mangiarotti, M. T. Bassi, G. Biamonti, and S. Riva. 1991.0. Alternative splicing in the human gene for the core protein hnRNP A1 generates another hnRNP protein. EMBO J. 9:1229-1235.

13. Chabot, B., M. Blanchette, I. Lapierre, and H. La Branche. 1997. An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1. Mol. Cell. Biol. 17:1776-1786.

14. Fu, X.- D. 1993. The superfamily of arginine/serine-rich family of splicing factors. RNA 1:663-680.

15. Manley, J. L., and R. Tacke. 1996. SR proteins and splicing control. Genes Dev. 10:1569-1579.

16. Zahler, A. M., W. S. Lane, J. A. Stolk, and M. B. Roth. 1992. S/R proteins -- a conserved family of pre-messenger-RNA splicing factors. Genes Dev. 6:837-847.

17. Screaton, G. R., J. F. Caceres, A. Mayeda, M. V. Bell, M. Plebanski, D. G. Jackson, J. I. Bell, and A. R. Krainer. 1995. Identification and characterization of three members of the human SR family of pre-mRNA splicing factors. EMBO J. 14:4336-4349.

18. Black D.L., 1995. Finding splice sites within a wilderness of RNA. RNA 1:763-71.

19. Hodges, D., and S. I. Bernstein. 1994. Genetic and biochemical analysis of alternative splicing. Adv. Genetics 31:207-281.

20. Reed, R. 1996. Initial splice-site recognition and pairing during pre-mRNA splicing. Curr. Opin. Genet. Dev. 6:215-220.

21. Tanabe, K. K., T. Nishi, and H. Saya. 1993. Novel variants of CD44 arising from alternative splicing: changes in the CD44 alternative splicing pattern of MCF-7 breast carcinoma cells treated with hyluronidase. Mol. Carcinogenesis 7:212-220.

22. Joensuu, H., P. J. Klemi, S. Toikkanen, and S. Jalkanen. 1993. Glycoprotein CD44 expression and its association with survival in breast cancer. Amer. J. Path. 143:867-874.

23. Dall, P., K.-H. Heider, H.-P. Sinn, P. Skroch-Angel, G. Adolf, M. Kaufmann, P. Herrlich, and H. Ponta. 1995. Comparison of immunocytochemistry and RT-PCR for detection of CD44v expression, a new prognostic factor in human breast cancer. Int. J. Cancer 60:471-477.

24. Iida, N., and L. Y. W. Bourguignon. 1995. New CD44 splice variants associated with human breast cancer. J. Cell. Physiol. 162:127-133.

25. Kaufmann, K.-H. Heider, H.-P. Sinn, G. von Minckwitz, H. Ponta, and P. Herrlich. 1995. CD44 variant exon epitopes in primary breast cancer and length of survival. Lancet. 345:615-619.

(9) APPENDICES

Figure Legends

Figure 1. CD44 alternative splicing. The diagram indicates the exon/intron structure of the murine CD44 gene. The ten variably-included exons v1-v10 are indicated in the second line. The standard form of CD44 contains none of these exons (i.e. exon 5 is directly spliced to exon 16). Three of the variable exons, v5–v7, have been associated with metastasis. Note that the human CD44 gene has a very similar structure to that of the murine gene, except that variable exon 1 does not exist. Exon sizes are indicated.

Figure 2. Expression of CD44 standard in breast tissue and breast tumors. Top panel) CD44 standard expression was detected in the indicated tissues by low cycle RT/PCR using primers specific for standard exons 5 and 16. The major band marked CD44 Std is a PCR product of 222 nucleotides. Larger products in tumor RNA result from inclusion of one or more variable exons in the interval between exons 5 and 16. Lane 1, 5 day pregnant gland, lane 2, lactating gland, lane 3, mature virgin gland, lane 4, hyperplasia TM2L (isolate 468/469), lane 5, preneoplasia TM40D (isolate 9016/9017), lanes 6-13, invasive breast adenocarcinomas, lane 14, liver metastasis derived from TM40D, lane 15, lung metastasis derived from TM40C. Bottom panel) RT/PCR analysis of hnRNP A1 mRNA splicing. The top band results from inclusion of alternative exon 7B, the bottom band from skipping this exon. PCR primers were directed against the exons flanking exon 7B.

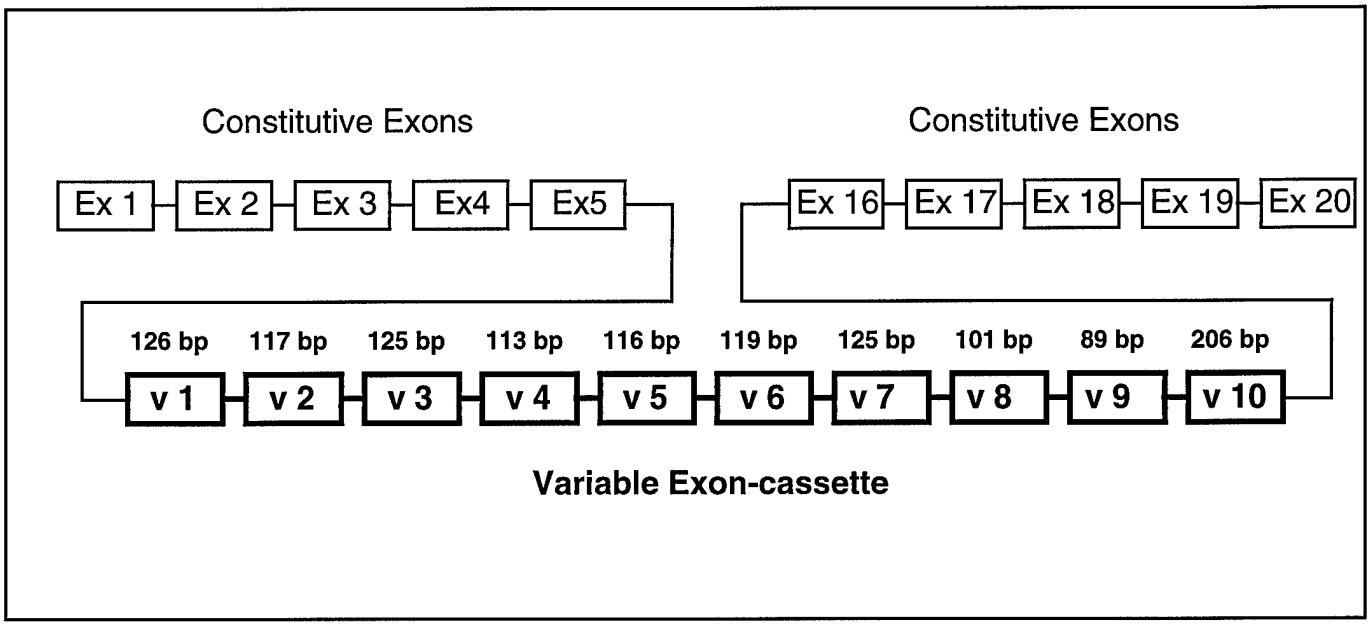
Figure 3. Detection of inclusion of variable exons v5, v6, and v7 by RT/PCR using exon-specific primers. Low cycle RT/PCR of indicated tissue RNA was performed using primers specific for exon 5 and variable exons v5 (A), v6 (B), or v7(C). Arrows mark the position of amplification bands from RNAs including only the targeted variable exon and no other upstream variable exons to give PCR amplification bands of 208 (v5), 210 (v6), or 209 (v7) nucleotides, respectively. Higher molecular weight bands resulting from inclusion of other upstream exons are also seen. Note that exons v1–v7 are all approximately 100 nucleotides in length so that the inclusion of each additional upstream intron results in a PCR product that is larger by this length. A and C) Lane 1, 5 day pregnant gland, lane 2, lactating gland, lane 3, D2 preneoplasia (D2 is a line that has been continuously grown in transplantation for over 30 years), lane 4, mature virgin gland, lane 5, hyperplasia TM2L (isolate 468/469), lane 6, preneoplasia TM40D (isolate 9016/9017), lanes 7-14, invasive breast adenocarcinomas, lane 15, liver metastasis derived from TM40D, lane 15, lung metastasis derived from TM40C. B) Lane 1, 5 day pregnant gland, lane 2, lactating gland, lane 3, mature virgin gland, lane 4, hyperplasia TM2L (isolate 468/469), lane 5, preneoplasia TM40D (isolate 9016/9017), lanes 6-13, invasive breast adenocarcinomas, lane 14, liver metastasis derived from TM40D, lane 15, lung metastasis derived from TM40C

Figure 4. High cycle RT/PCR amplification of variable exons 5 and 7 from normal and preneoplastic tissue. To better examine the low amounts of RNA produced to include variable exons v5 and v7, amplifications using 35 cycles were performed.

Arrows mark the amplification band resulting from inclusion of only the targeted variable exon. Lane 1, pregnant gland; lane 2, mature virgin gland; lane 3, hyperplasia TM2L; lane 4, preneoplasia TM40D.

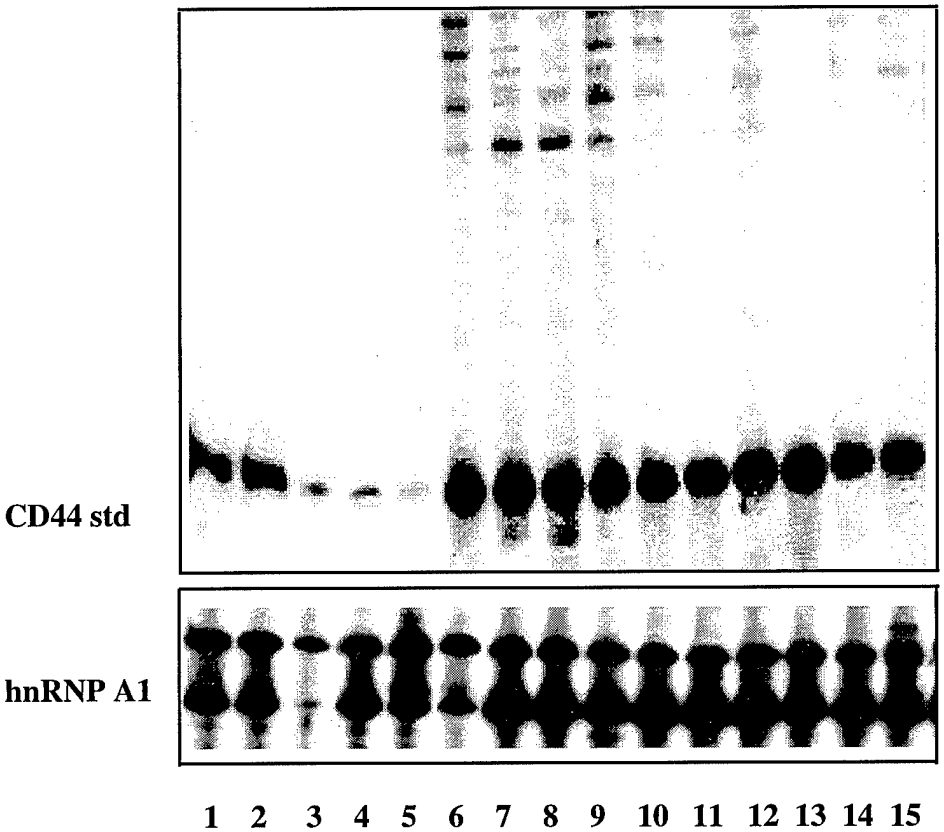
Figure 5. Detection of inclusion of variable exon v3. Low cycle RT/PCR of indicated tissue RNA was performed using primers specific for exon 5 and variable exons v3. The arrow marks the position of an amplification band from RNA including only the targeted variable exon and no other upstream variable exons. Higher molecular weight bands resulting from inclusion of other upstream exons are also seen. Lane 1, 5 day pregnant gland, lane 2, lactating gland, lane 3, D2 preneoplasia (D2 is a line that has been continuously grown in transplantation for over 30 years), lane 4, mature virgin gland, lane 5, hyperplasia TM2L (isolate 468/469), lane 6, preneoplasia TM40D (isolate 9016/9017), lanes 7-14, invasive breast adenocarcinomas, lane 15, liver metastasis derived from TM40D, lane 15, lung metastasis derived from TM40C.

Figure 6. Expression of SR protein splicing factors in normal breast tissue and breast tumors. SR proteins from the indicated tissues were detected on Western blots of total tissue protein using the monoclonal antibody mAb104 specific for a phosphorylated epitope present on the major SR proteins, SRp75, SRp55, SRp40, SRp30a, SRp30b, and SRp20. Lane 1, 5 day pregnant gland; lane 2, 15 day pregnant gland; lane 3, 1 day lactating gland; lanes 4-6, invasive adenocarcinomas of the breast.



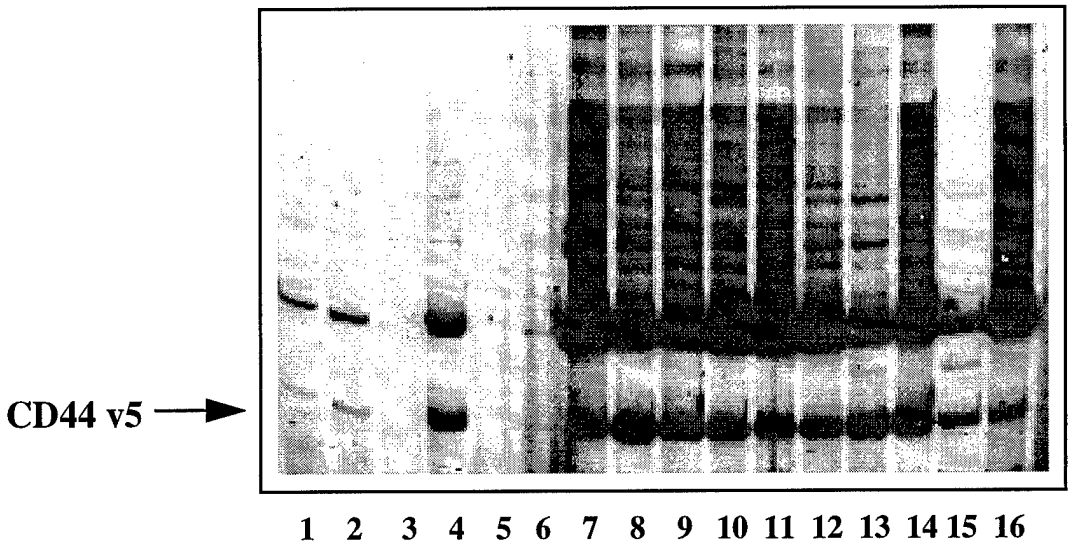
Susan M. Berget
DAMD17-96-1-6084

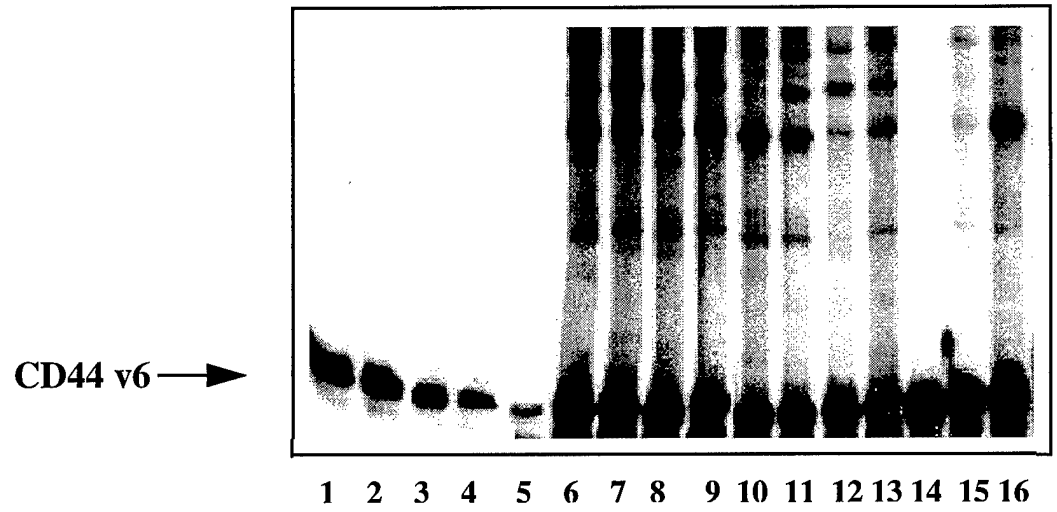
Figure 1



Susan M. Berget
DAMD17-96-1-6084

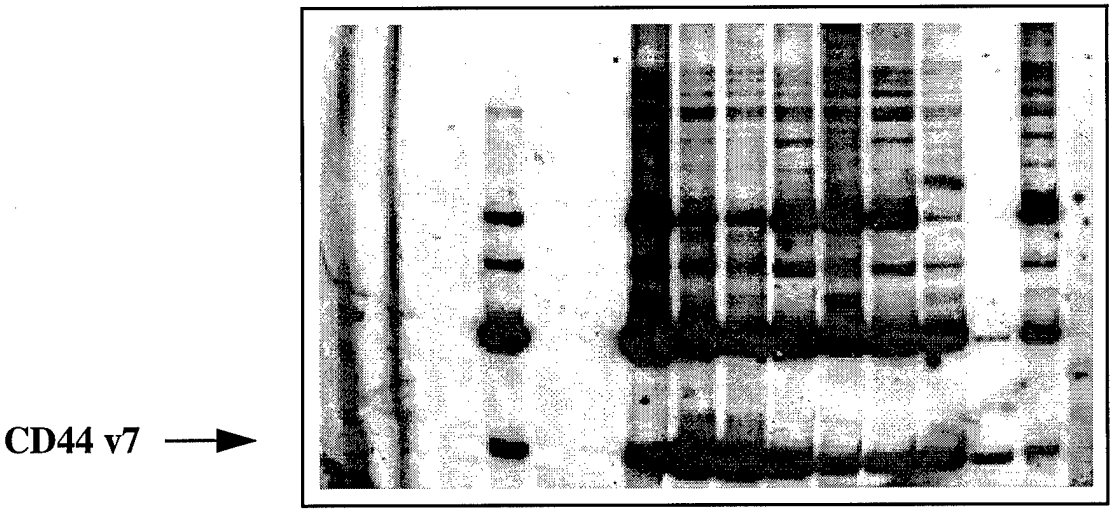
Figure 2





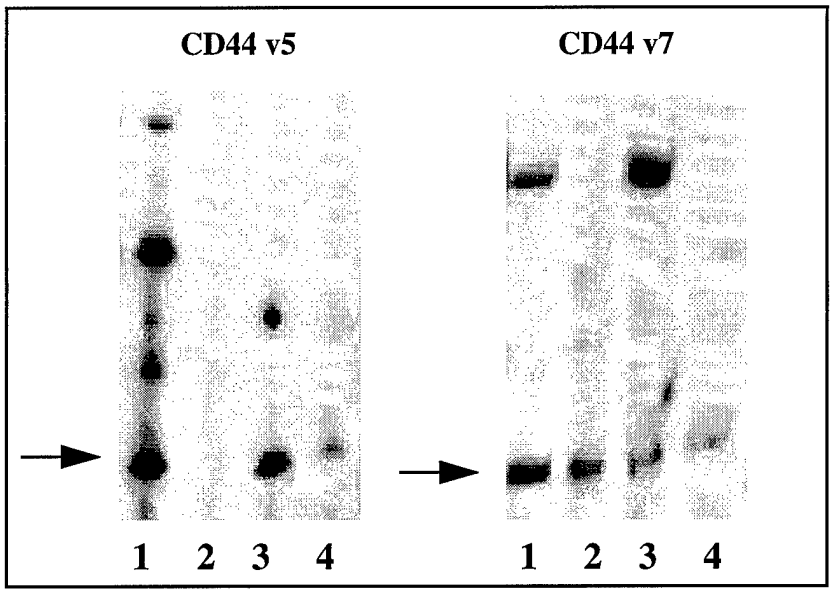
Susan M. Berget
DAMD17-96-1-6084

Figure 3A & 3B



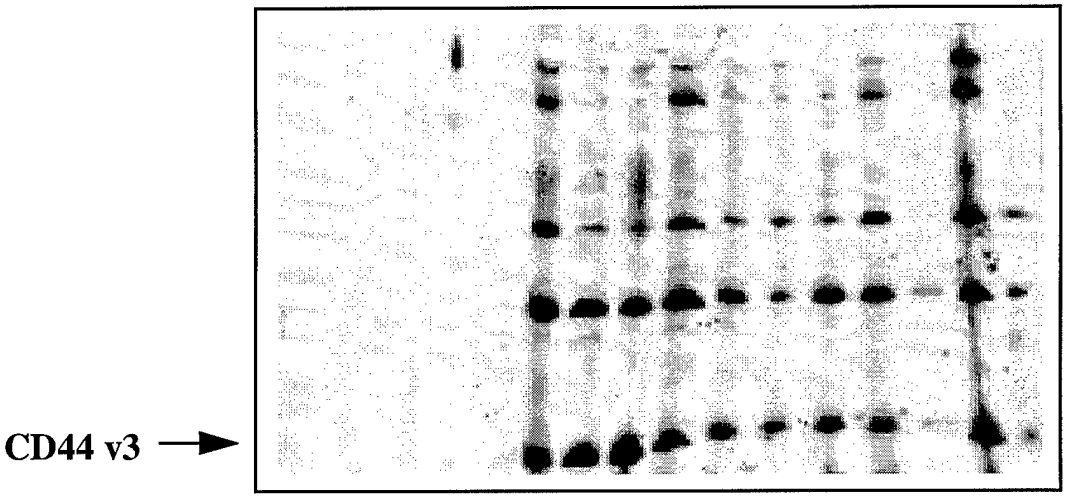
Susan M. Berget
DAMD17-96-1-6084

Figure 3C



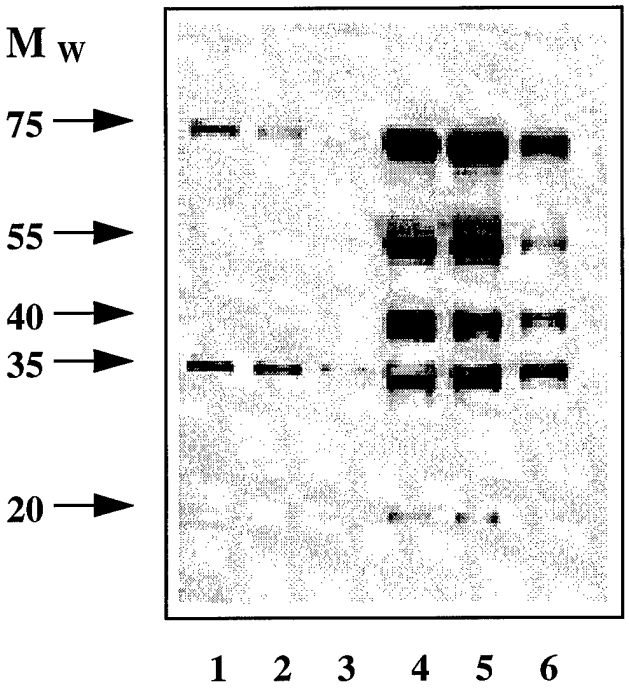
Susan M. Berget
DAMD17-6-1-6084

Figure 4



Susan M. Berget
DAMD17-96-1-6084

Figure 5


M w
75
55
40
35
20
1 2 3 4 5 6